CHAPTER 2
Baseband Modulation and Demodulation

2.1 Line Codes:

The output of the PCM coder is a bit stream. It is now required to transmit this bit stream over a communication channel. It may be done over a copper wire, coaxial cable, an electromagnetic wave through the atmosphere or an optical signal in an optical fiber. The modulator is a device that turns the digital bit stream into a waveform that can be detected. The binary digits are usually represented by electrical pulses which can be transmitted over the baseband channel i.e. without modulation. The pulse has one dc value for digit 1, and another dc value for digit 0. This is called baseband transmission. At the receiver, a determination must be made as to whether we are receiving 0 or 1. If digit 1 is represented by $5V$ and digit 0 is represented by OV, then the presence of a voltage at the receiver indicates digit 1 and the absence of the voltage 1 indicates level 0. It will be shown that the likelihood of correctly detecting the presence of a pulse is a function of the received pulse energy (area under the pulse). Thus, there is an advantage in making the pulse width τ as wide as possible. If we increase the pulse width to the maximum possible (bit time slot $\tau = T_b$) we have the a waveform shown (Fig. 2.1).

Rather than describe this waveform as a sequence of present or absent pulses (unipolar) (Fig. 2.1b), we can describe it as a sequence of transitions between two levels (bipolar). When the waveform occupies the upper voltage level, it represents binary 1. When it occupies the lower voltage level it represents binary 0.

There are several types of PCM waveforms, usually called line codes in telephony applications, and are divided into 4 main groups .

1. Non return to zero (NRZ)
2. Return to zero (RZ)
3. Phase encoded (PE)
4. Multilevel binary (MB)

The NRZ group is the most commonly used. Binary 1 is represented by one voltage level, and binary 0 is represented by another voltage level. There is a change in level whenever the data change from 1 to 0 or from 0 to 1. With NRZ-M 1 is represented by a change in level, while 0 is represented by no change in level. This is called differential encoding, and is used in magnetic recording. The NRZ-S is the complement. In unipolar RZ, 1 is represented by half bit wide pulse and 0 is represented by the absence of pulse.

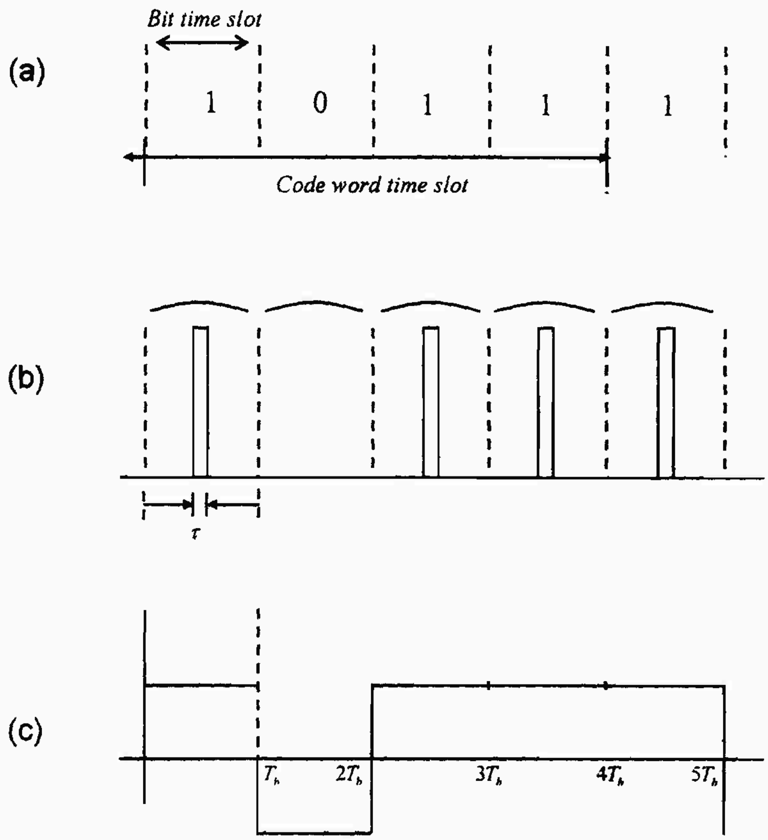


Fig. (2.1) Waveform representation of binary digits

a) PCM bit sequence

b) unipolar pulse representation with $\tau < T_b$

c) bipolar pulse representation with $\tau = T_b$

In bipolar RZ, 1 , 0 are represented by opposite level pulses, each half bit wide. In RZ-AMI (alternate mark inversion), $1_s'$ are represented by equal amplitude alternating pulses, $0_s'$ are represented by the absence of pulses. The phase encoded group consists of $bi-\varphi-L$ (bi-phase-level) or Manchester coding, $bi-\phi-M$ (bi-phase-mark), $bi-\phi-s$ (bi-phase-space) and delay modulation (DM) or Miller coding.

In $bi-\phi-L$, 1 is represented by a half bit wide pulse positioned during the first half of the bit interval, and 0 is represented by a half bit wide pulse positioned during the second half of the bit interval. In $bi-\varphi-M$, a transition occurs at the beginning of every bit interval. A 1 is represented by a second transition one half bit interval later, and 0 is represented by no second transition. In $bi-\varphi-S$, a transition also occurs at the beginning of every bit interval. A 1 is represented by no second transition, and 0 is represented by a second transition one half bit interval later. In delay modulation (Miller Coding), a 1 is represented by a transition at the midpoint of the bit interval. A 0 is represented by no transition, unless it is followed by another 0. In this case, a transition is placed at the end of the bit interval of the first zero. Multi binary waveforms are 3 levels instead of 2, to encode the binary data. Bipolar RZ and RZ-AMI belong to this group.

This group also contains dicode (duobinary) formats. In dicode NRZ, the one-to-zero or zero-to-one data transition changes the pulse polarity. Without a data transition, the zero level is sent (1 is $-V$ and 0 is $+V$). With dicode RZ, the one-to-zero or zero-to-one transition produces a half duration polarity change, otherwise a zero level is sent.

In choosing a PCM waveform for a particular application, the following parameters have to be considered

1. DC component: Eliminating the dc energy from the signal power spectrum enables the system to be ac coupled. Magnetic recording systems or transformer-coupled systems are not sensitive to dc or low frequency.
2. Self clocking: Symbol or bit synchronization is required for any digital communication system. Some PC coding schemes have inherent synchronizing (clocking) feature. The Manchester code – e.g. – has a transition at the middle of every bit interval.
3. Error detection: Some schemes provide means for error detection such as RZ AMI (Fig. 2.3).
4. Bandwidth compression: Some schemes such as multilevel codes increase the efficiency of bandwidth utilization by allowing a reduction in the required bandwidth for a given data rate. Thus, there is more information transmitted per unit bandwidth. Consider NRZ-L and bipolar RZ (Fig. 2.4). In NRZ-L, the waveform is twice as long as in bipolar RZ, so the bandwidth of NRZ-L is half that of bipolar RZ.

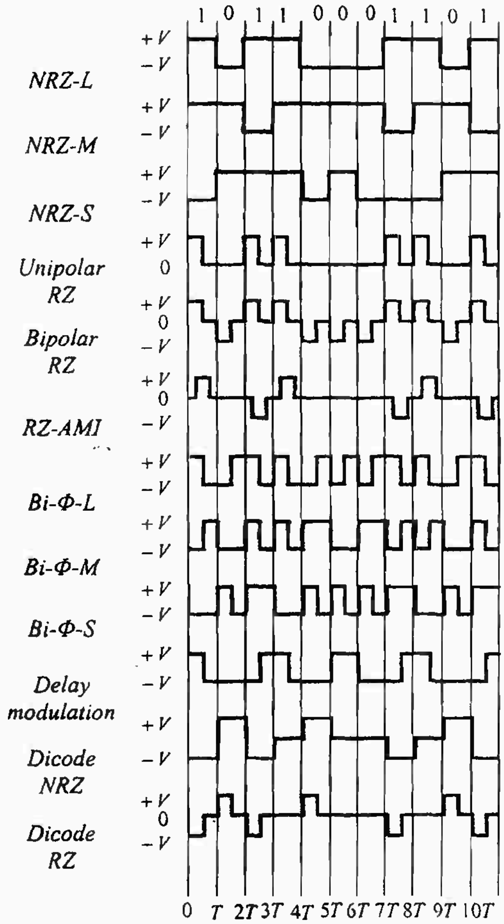


Fig. (2.2) PCM waveforms (Line codes)

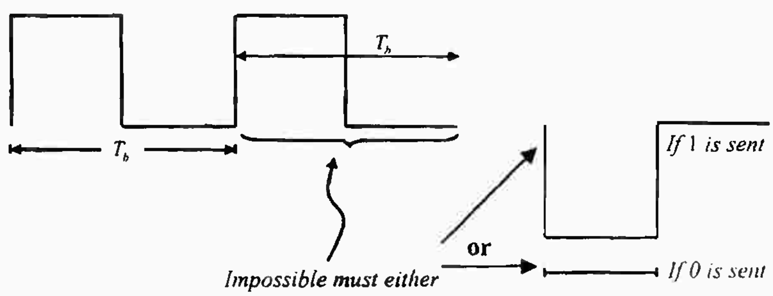


Fig. (2.3) RZ-AMI detecting an error

In RZ-AMI (1's are detected by alternating pulses and 0's are spaces (no pulses)

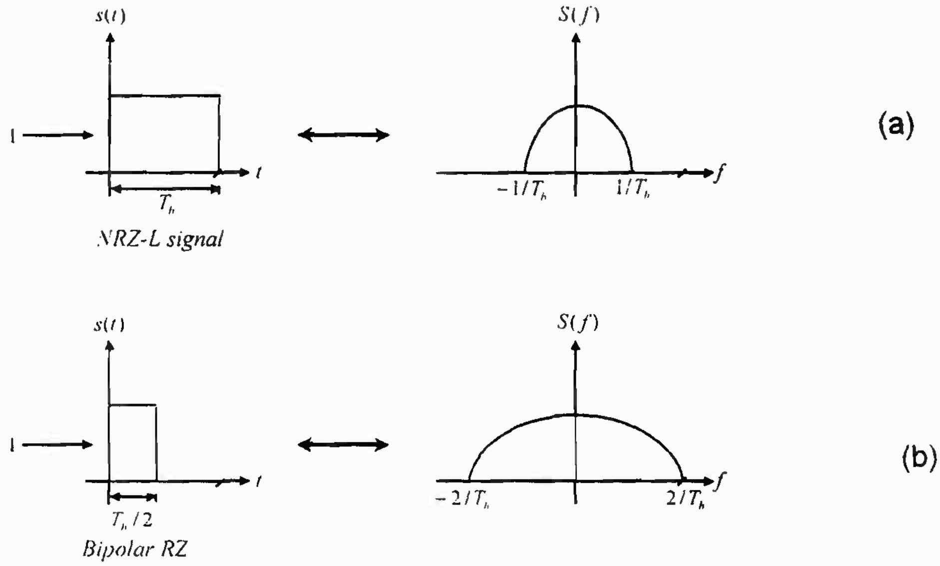


Fig. (2.4) Bandwidth compression
a) NRZ-L *b) Bipolar RZ*

5. Noise immunity: Some of the schemes are more immune than others to noise. For example, the NRZ waveforms have better error performance than uniform RZ waveforms.

An important figure of merit is the normalized spectral bandwidth which is BT_s, where B is bandwidth and T_s the duration of the symbol pulse. Thus BT_s is often referred to as time bandwidth product. Since the symbol rate R_s is $1/Ts$, the normalized bandwidth is B/R_s. It describes how efficiently the transmission bandwidth is being utilized. A waveform that requires less than $1\,Hz$ for sending 1 symbol/s is bandwidth efficient.

Another parameter is bandwidth efficiency R_b/B where R_b is the bit rate. It measures how much data (bits per second) not symbol rate-are transmitted per Hz.

Ex. 2.1

Determine the number of bits per symbol (PCM word size) for quantization distortion error $\epsilon \leq p\,V_{pp}$ where p is a fraction of V_{pp}.

Solution

$$|\epsilon|_{\max} = \frac{q}{2} = \frac{V_{pp}}{2(L-1)} \approx \frac{V_{pp}}{2L} \qquad (2-1)$$

where q is the quantum interval, and L is the number of quantization levels. Thus,

$$\frac{V_{pp}}{2L} \le p\, V_{pp} \qquad (2-2)$$

For PCM,

$$L = 2^n \qquad (2-3)$$

where n is the number of bits per symbol. From eqns (2 – 2) and (2 – 3),

$$2^n = L \ge \frac{1}{2p}$$

$$n \ge \log_2\left(\frac{1}{2p}\right) \qquad (2-4)$$

2.2 M-ary Pulse Modulation:

There are 3 basic ways to modulate information onto a sequence of pulses. We may change the amplitude, giving rise to pulse amplitude modulation (PAM), or change the position (PPM) or duration (PDM) / width (PMW). In analog PAM the information samples are modulated onto a pulse train without any quantization. Digital PAM means that the information is first quantized, yielding symbols from an M-ary alphabet set, and then modulated onto pulses. This is called M-ary pulse modulation. One of M allowable levels is assigned to each of the M possible symbols. In PCM, we have binary waveforms, a special case of M-ary PAM with $M = 2$. For k bit group, $M = 2^k$. Thus, in binary PCM, $k = 1$. We need one bit to describe the state or the level of the signal.

Multilevel signaling-such as M-ary PAM-has the advantage of reducing the bandwidth required. Consider a bit stream with data rate R_b $bits/s$. Instead of transmitting a pulse for every bit we partition the data into k bit groups according to $M = 2^k$ level pulses. Each pulse can now represent, the k bit symbol rate is $R_b / k = R_S$ Sym/s. The price we have to pay for bandwidth or symbol rate reduction is two fold. The receiver must distinguish between each level of each pulse instead of distinguishing between just two levels? Also, for equal average power in binary and octal pulses, the binary PCM will have more signal energy per level than an 8 level system. In PCM we call the number of level $L = 2^n$ where n is the word size. In PAM, i.e. $M = 2^k$, where k is the symbol size and M is the number of level pulses. We must draw here a comparison between PCM and PAM. The transmission bandwidth required for PCM is very large. To reduce this bandwidth we use multilevel signaling. We partition the data into the bit groups, using $M = 2^k$ level pulses per transmission. With such multilevel signaling in M-ary PAM each pulse represents a k bit symbol in a bit stream moving at a rate of R_b / k (symbols/s) where R_b is the bit rate. Thus the transmission bandwidth is reduced for the same data rate, or the data rate may be increased for the same bandwidth. What is the price to be paid? The receiver must distinguish between the possible levels of each pulse. The transmission of an 8 level pulse compared with a 2 level pulse requires a greater energy for the same detection performance.

For equal average power in the binary and octal pulses, it is easier to detect the binary pulses because the detector has more signal energy per level for making a binary decision than in 8 level system. If we choose PCM for ease for detection we have to settle for an increased bandwidth (3 times that in octal transmission). This is the trade off to be made.

Ex. 2.2

An analog signal of maximum frequency $3kHz$ is to be transmitted over an M-ary PAM system where the number of levels $M = 16$. The quantization distortion is not to exceed $\pm 1\%$ of the peak to peak signal

a) Find the PCM word size
b) Find the minimum sampling rate, bit rate and symbol rate
c) For a transmission bandwidth $12kHz$, determine the bandwidth efficiency.

Solution

a) From eqn (2 – 4), $n \geq \log_2 \frac{1}{0.02} = \log_2 50 = 5.6$. Therefore, $n = 6$ to meet the distortion condition i.e, the PCM word size is 6 bits/sample.
b) Using the Nyquist sampling criterion, the minimum rate $f_s = 2f_M = 6k$ samples/s. The bit rate for $n = 6$ is $R_b = 36k\ b/s$. Since multilevel pulses are to be used with $M = 2^k = 16$, $k = 4$ bits/symbol. The bit stream will be grouped into 4 bits to form a 16 level PAM digit. The resulting symbol rate $R_s = R_b / k = \frac{36}{4} = 9$ symbols/s.
c) The bandwidth efficiency is described by data throughput per Hz. Since $R_b = 36k\, b/s$ and given B=12 kHz , then $R_b / B = 3$ bits/s/Hz.

2.3 Channel Bandwidth Limitations:

The output signal from an ideal channel may have some time delay compared with the input, but must have the same shape as the input. Thus, for a distortionless channel, the output signal $y(t)$ is given by

$$y(t) = K\,x(t - t_o) \qquad (2-5)$$

where K is a scalar gain factor which is equally valid for all frequency components of the Fourier transform of the time signal, and t_o is a time delay for the time signal. Taking Fourier transform of both sides

$$Y(f) = K\,X(f)\,e^{-j2\pi f t_o} \qquad (2-6)$$

Since

$$Y(f) = H(f)\,X(f) \qquad (2-7)$$

where $H(f)$ is the system tranfer function, it must be given by

$$H(f) = K\,e^{-j2\pi f t_o} \qquad (2-8)$$

Therefore, to achieve ideal distortionless transmission, the overall system response must have a constant magnitude response, and its phase shift must be linear with frequency. All the frequency components must have the same time delay in order to add up correctly. For t_o to be constant, the phase shift ϕ must be

$$\phi = 2\pi f t_o \qquad (2-9)$$

The linearity of the phase shift with frequency – called envelope delay or group delay – is used as a measure the delay distortion, and is given by

$$t_o(f) = -\frac{1}{2\pi}\frac{d\phi(f)}{df} \qquad (2-10)$$

Thus, for distortionless transmission, $t_o(f)$ must be constant and $\phi(f)$ is a linear function of frequency. Phase or amplitude distortion may be introduced into the system. Equalizing networks are often used for the correction of such distortion.

The ideal network or channel of eqn (2 – 8) cannot be realized because it entails a constant response independent of frequency for all values of frequency from 0 to ∞, i.e., infinite bandwidth. In practice $H(f)$ can remain constant only for part of the spectrum called channel bandwidth B_c. An approximation to the ideal infinite bandwidth network may be one of three types. A band pass filter (BPF) has a truncated spectrum which passes without distortion all frequency components between f_l and f_u, where f_l is the lower cut off frequency and f_u is the upper cut off frequency. (Fig. 2.5a). The pass band is defined as frequencies lying between f_l and f_u. Outside this pass band, the network is assumed to have zero response.

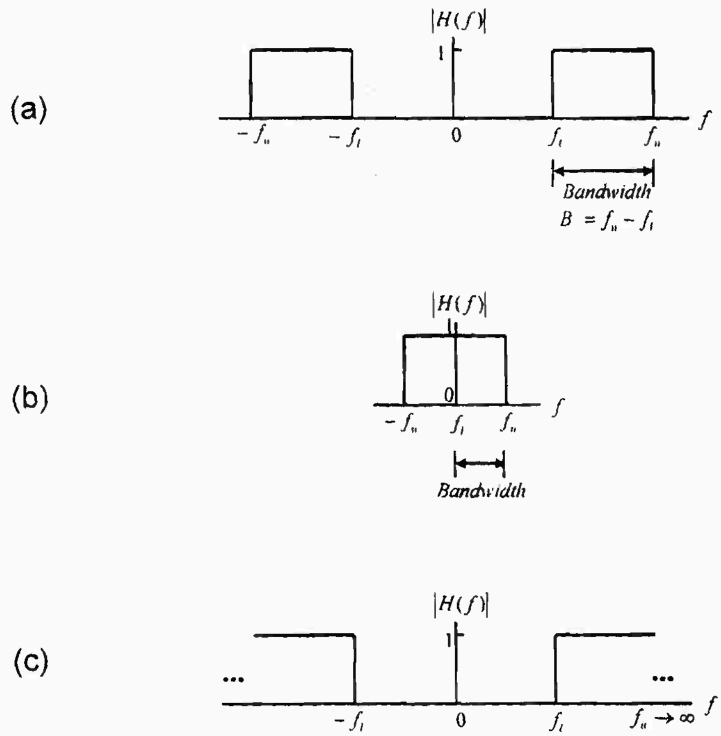


Fig. (2.5) Ideal filter transform function

a) ideal BPF *b) ideal LPF* *c) ideal HPF*

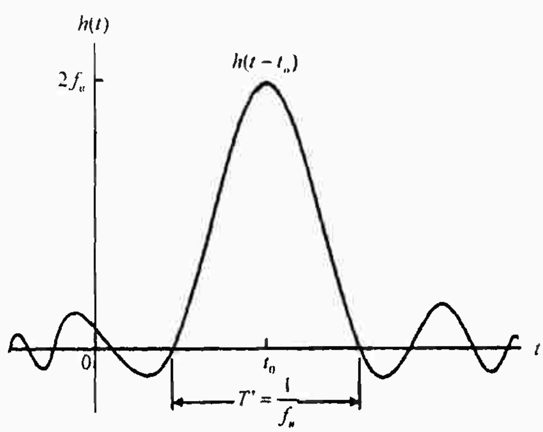


Fig. (2.6) Impulse response of an ideal LPF

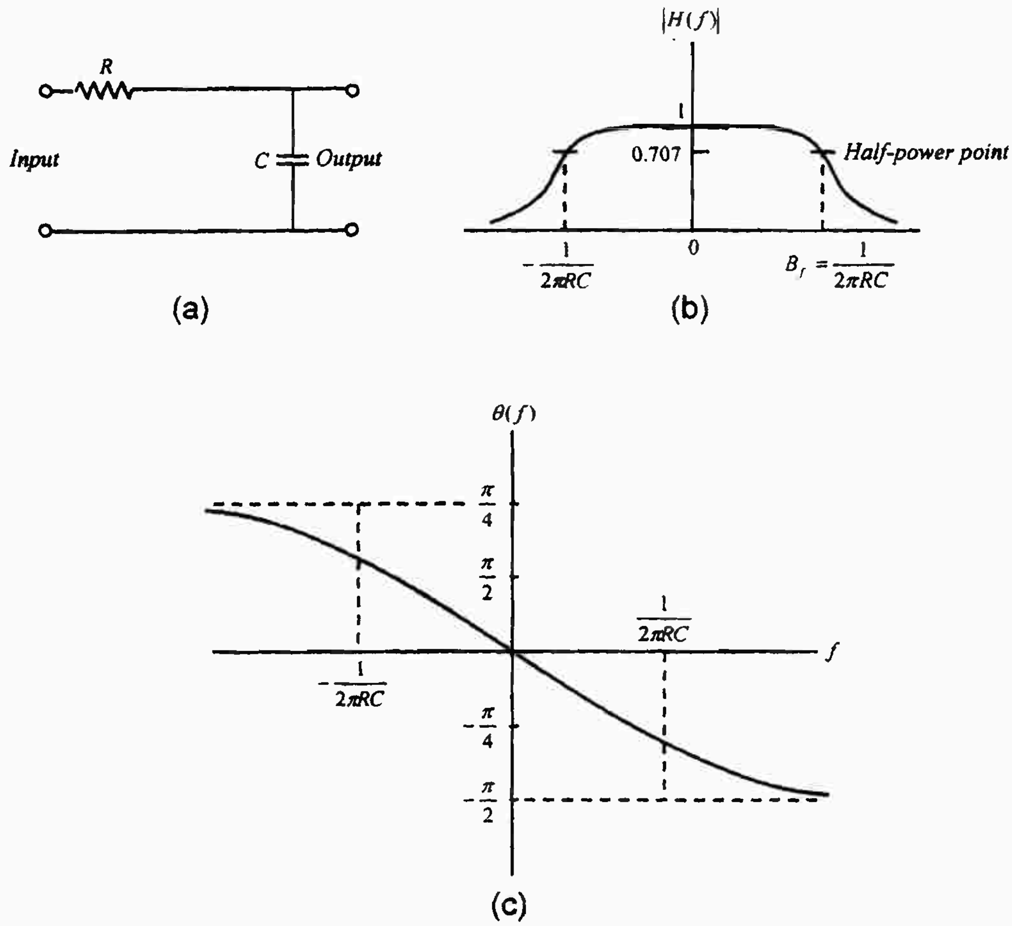


Fig. (2.7) RC filter

a) circuit *b) magnitude characteristic* *c) phase characteristic*

This pass band is the channel (or network) bandwidth $B_c = f_u - f_l$. If $f_l = 0$ and f_u has a finite value the filter is called low pass filter (LPF) (Fig. 2.6b). If f_l has a non zero value and $f_u = \infty$, the filter is called high pass filter (HPF) (Fig. 2.6c). Following eqn. (2 – 8), and letting $K = 1$ for the ideal LPF with $B_c = f_u$,

$$|H(f)| = \begin{cases} 1 & |f| < f_u \\ 0 & |f| \geq f_u \end{cases} \qquad (2-11)$$

The impulse response $h(t)$ of the ideal LPF is given by

$$h(t) = \mathcal{F}^{-1}[H(f)] = \int_{-\infty}^{\infty} H(f)\, e^{j2\pi f t} df \qquad (2-12)$$

$$= \int_{-f_u}^{f_u} e^{-j2\pi f t_o}\, e^{j2\pi f t} df$$

$$= 2f_u \frac{\sin 2\pi f_u (t - t_o)}{2\pi f_u (t - t_o)}$$

$$= 2f_u \operatorname{sinc} 2f_u (t - t_o) \qquad (2-13)$$

where sin cx is defined as

$$\operatorname{sinc}\ x = \frac{\sin \pi x}{\pi x} \qquad (2-14)$$

$$x = 2f_u (t - t_0) \qquad (2-15)$$

Hence, $h(t)$ is shown (Fig. 2.6). It is seen that it has non zero output prior to the application of the input at $t = 0$. Thus, an ideal filter cannot be realized. The simplest practical approximation of a LPF is an RC filter shown (Fig. 2.7).

Let us consider now a band limited signal to be inputted to a band limited system or channel. If the input signal has a narrow spectrum and the channel has a wide bandwidth, the output which is the product of the two spectra will be limited by smaller of the two bandwidths. Thus, in Fig. (2.8), the output signal is constrained by the input spectrum (case I). While in case II, the output signal is constrained by the filter (channel) bandwidth. In this case the output signal will be a distorted version of the input signal due to filtering. The filtering effect can be viewed in the time domain. Assume an input pulse of height V_m and pulse width τ is applied to a LPF. The output is given by

$$y(t) = V_m\left(1 - e^{-t/CR}\right) \qquad o \le t \le \tau \qquad (2-16)$$

$$= V_m'\ e^{-(t-\tau)/CR} \qquad t > \tau \qquad (2-17)$$

where $V_m' = V_m\left(1 - e^{-\tau/CR}\right)$ (2 – 18)

Let us define the pulse bandwidth as

$$B_p = \frac{1}{\tau} \qquad (2-19)$$

The RC filter bandwidth is

$$B_f = \frac{1}{2\pi RC} \qquad (2-20)$$

The ideal input pulse $x(t)$ and its magnitude spectrum $X(f)$ are shown (Fig. 2.9)

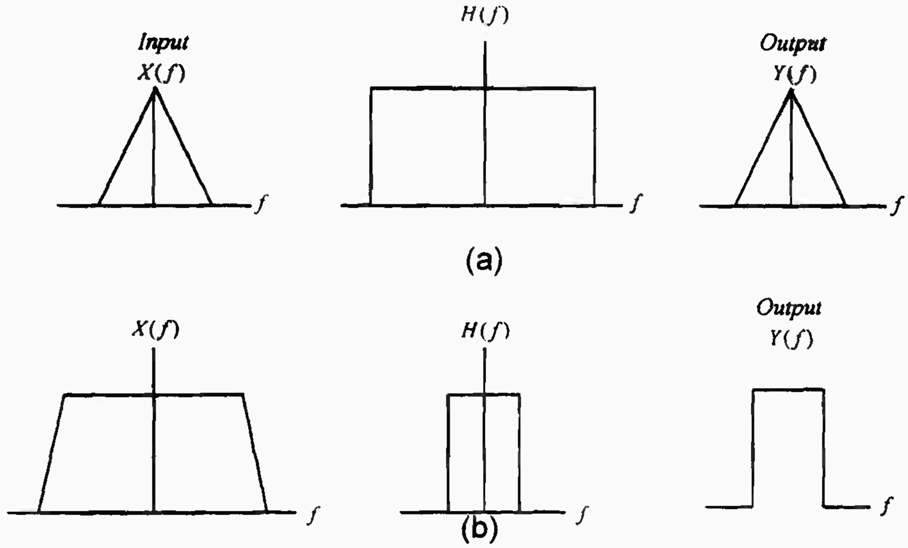


Fig. (2.8) Output characteristics *Vs* input and circuit characteristics
a) output is restricted by the input characteristics . (case I)
b) output is restricted by the circuit characteristics. (case II)

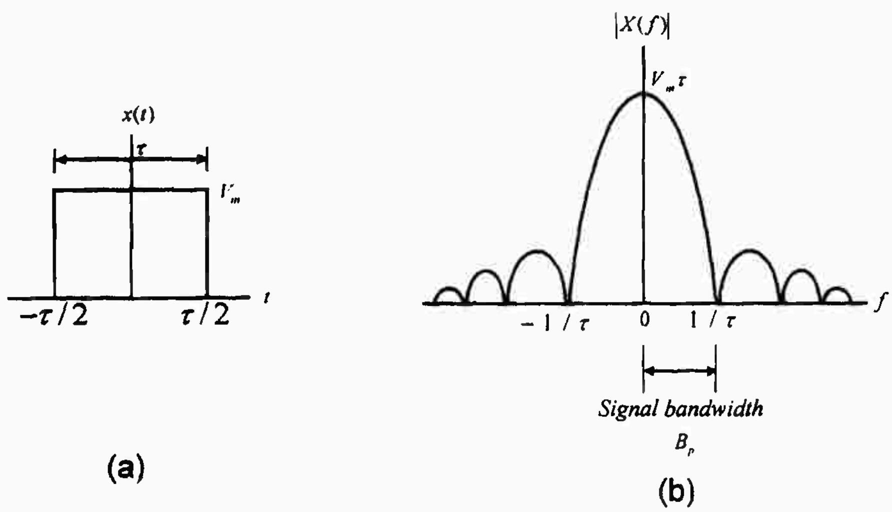


Fig.(2.9) Ideal pulse and its magnitude spectrum
a) ideal pulse *b) magnitude spectrum*

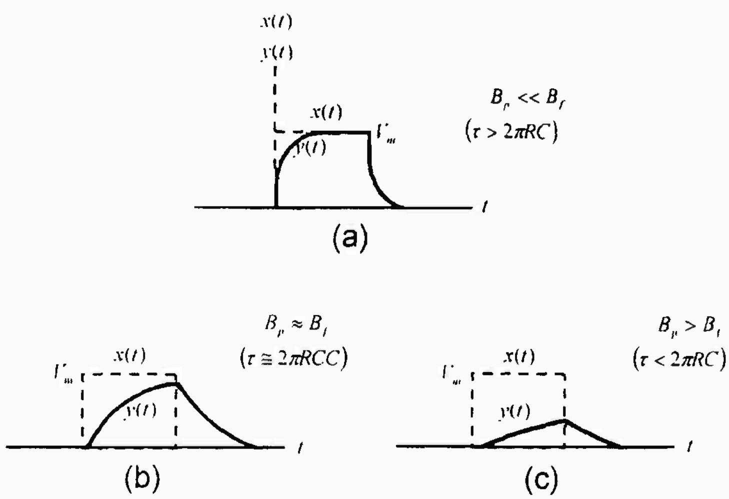


Fig. (2.10) Filtering an ideal pulse

a) $B_p < B_f$ *b)* $B_p = B_f$ *c)* $B_p > B_f$

For an RC filter, we may consider three cases. Case 1 is when $B_p < B_f$, the output $y(t)$ is a good approximation of the input $x(t)$ (Fig 2.10a). In case $B_p = B_f$ the output is a distorted replica of the input (Fig. 2.10b). Case 3 is when $B_p < B_f$, the pulse is hardly perceptible from $y(t)$ (Fig. 2.10c). In digital communication, case 2 is sufficient, since we need only recognize the presence of the pulse. The quality of the pulse shape is not critical in digital transmission.

If a signal is transmitted as baseband, its bandwidth is f_M, which is the highest frequency of the signal. If it is transmitted as double side band modulated (DSB) signal its DSB bandwidth is $2f_m$, where the carrier frequency $f_c >> f_M$ (Fig. 2.11).

In most communication systems, we are dealing with band limited channels. This means that the signal power is not allowed outside a defined band. The band limiting of the signal may come about from the signal source, the electronics involved, or from the channel. In any case, a strictly band limited signal is not physically realizable, since a signal limited in the frequency domain is infinite in the time domain and vice versa.

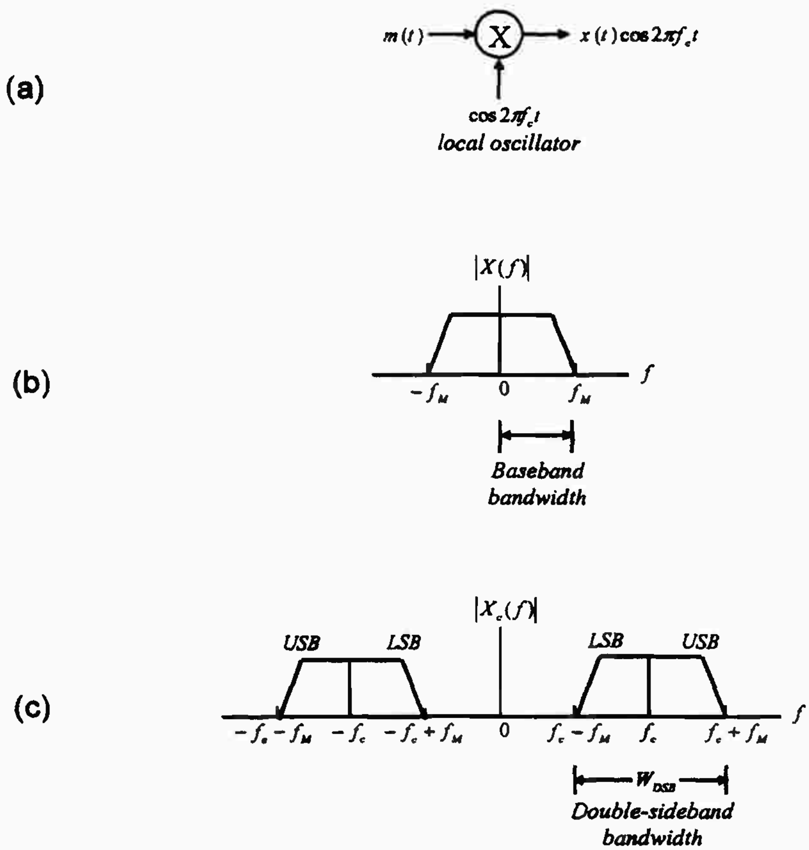
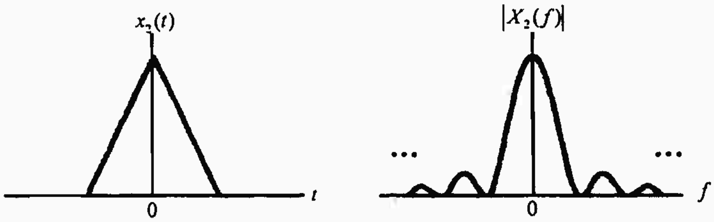


Fig. (2.11) Comparison of baseband and double side band (DSB) spectra
a) heterodyning b) baseband spectrum c)(DSB) spectrum

In fact, the uncertainty principle dictates that the product of the time width and the frequency bandwidth is constant. In Fig. (2.7) $\tau f_u = 1$ which means if τ is small, f_u is large and vice versa. Fig. (2.12a) shows a signal limited in frequency. It has an extensive duration in time, whereas Fig. (2.12b) shows a duration - limited signal whose spectrum is extensive in frequency. We can always define an effective time duration and an effective frequency bandwidth, such that the uncertainty principle holds.

(a) (b)

Fig. (2.12) Strictly band limited signal
a) a band limited signal in the time domain
b) a band limited signal in the frequency domain



(a) (b)

Fig. (2.13) Strictly duration-limited signal
a) a time limited signal in the time domain
b) a time limited signal in the frequency domain

Ex. 2.3

Show that a bandpass, band limited signal $x(t)$ of bandwidth B has infinite time duration.

Solution

For signal $x(t)$ with Fourier transform $X(f)$ limited to the band of frequencies centered at $\pm f_c$ of width $2B$, we may express $X(f)$ in terms of an ideal filter transfer function $H(f)$ (Fig. 2.15) as

$$X(f) = X'(f)\, H(f) \qquad (2-21)$$

where $X'(f)$ is the Fourier transform of $x'(t)$ which in the time function before filtering

$$H(f) = rect\left(\frac{f - f_c}{2B}\right) + rect\left(\frac{f + f_c}{2B}\right) \qquad (2-22)$$

where

$$rect\left(\frac{f}{2B}\right)=\begin{cases}1 & -B<f<B\\ 0 & |f|>B\end{cases} \qquad (2-23)$$

$$X(f)=\begin{cases}X'(f) & (f_c-B)\le|f_c|\le f_c+B\\ 0 & \text{otherwise}\end{cases} \qquad (2-24)$$

Multiplication in the frequency domain (eqn 2 – 21) transforms to convolution in the time domain

$$x(t)=x'(t)*h(t) \qquad (2-25)$$

where $h(t)$ - the inverse Fourier transform of $H(f)$ - can be written as

$$h(t)=2B\sin 2Bt\cos 2\pi f_c t \qquad (2-26)$$

We note from Fig. (2.14) that $h(t)$ has infinite duration, and thus $x(t)$ also has infinite duration and, therefore, is not realizable

2.4 Intersymbol Interference:

The channel is always band limited. A band limited channel disperses or spreads a pulse waveform passing through it, when the channel bandwidth is much greater than the pulse bandwidth, the spreading of the pulse will be slight. When the channel bandwidth is close to the signal bandwidth, the spreading will exceed the symbol duration, and cause signal pulses to overlap. This overlapping is called intersymbol interference (ISI) This causes system degradation, i.e., high error rates.

We represent the communication channel by an overall transfer function $H(f)$ given by

$$H(f)=H_t(f)\,H_c(f)\,H_r(f) \qquad (2-27)$$

where $H_t(f)$ is the transfer function of the filter representing the transmitter circuitry, $H_c(f)$ the transfer function of the channel and $H_r(f)$ the transfer function of the receiver circuitry. At the transmitter, the bandwidth of the sample pulses is limited by the circuit bandwidth constraint. The channel introduces distortion (amplitude, frequency, phase) due to distributed reactance which affects the baseband signal. It is a good approximation to regard the incoming pulses as delta shaped impulses seprated by symbol duration T_s . This applies to RZ pulses. It also applies to NRZ pulses considering the effect of receiver sampling as a train of delta functions. The receiving filter is usually designed to make up for such distortion as much as possible, and hence is called an equalizing filter.

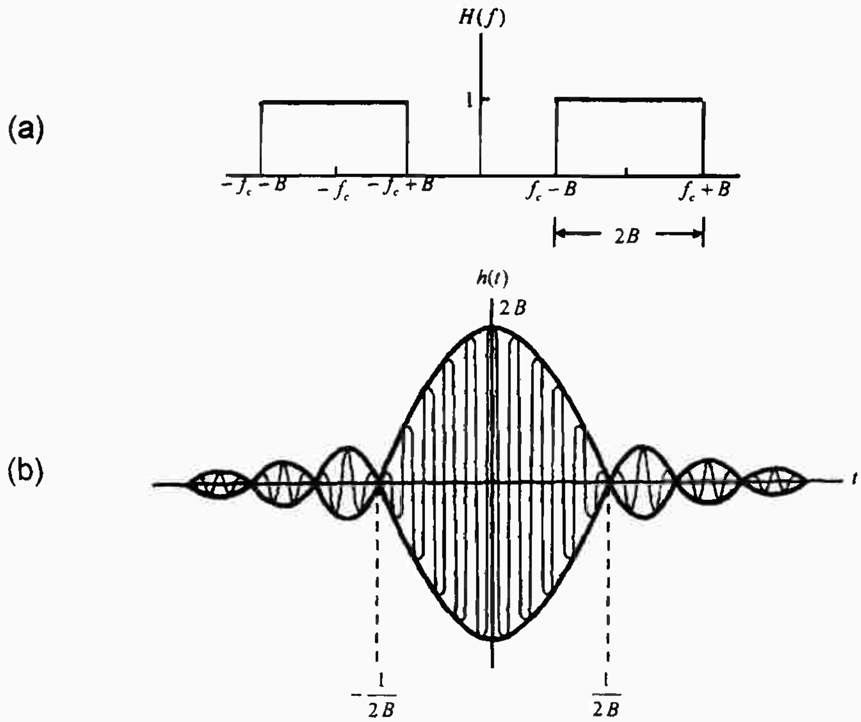


Fig. (2.14) Bandpass limited signal

a) ideal band pass filter (BPF)   b) ideal BPF impulse response

In a binary system with a common PCM waveform (such as NRZ-L), the detector makes a symbol decision by comparing the sample of the received pulse to some threshold (positive for 1, and negative for 0). Due to the effects of system filtering, the received pulses can overlap one another (Fig. 2.15). The tail of a pulse can smear into adjacent symbol intervals, thereby interfering with the detection process and degrading the error performance. This is intersymbol interference (ISI) even without noise. Usually, $H_c(f)$ is specified, hence the designer is concerned with the determination of $H_t(f)$ and $H_r(f)$ to minimize ISI at the output $H_r(f)$. Nyquist investigated the problem of specifying a received pulse shape for zero ISI. He showed that the theoretical minimum system bandwidth needed to detect R_s symbols $/s$ without ISI is $R_s/2\ Hz$. To see this assume $H(f)$ to be rectangular with single sided bandwidth $1/2T_s$ (Fig. 2.16). For baseband systems, this ideal.

Nyquist filter has an impulse response in the form $h(t) = \text{sinc}(t/T_s)$, which is the inverse Fourier transform of rectangular $H(f)$ of single sided bandwidth $1/2T_s$. This $\sin c$ shaped pulse is called ideal Nyquist pulse. Its multiple lobes comprise a mainlobe and side lobes called tails that are infinitely long. Nyquist established that if each pulse of a received sequence is of the form $\sin c(t/T_s)$. The pulses can be detected without ISI, if there are two successive pulses $h(t)$ and $h(t-T_s)$ (Fig. 2.16), even though $h(t)$ has long tails. The figure shows a tail passing through zero amplitude at the instant $t = T_s$ when $h(t-T_s)$ is to be sampled. Thus, no ISI degradation is introduced. For baseband systems the bandwidth required to detect $1/T_s$ such pulses (or R_s symbols per second) is $1/2T_s$. Thus a system with bandwidth $B = 1/2T_s = R_s/2\ Hz$ can support a max transmission rate of $2B = 1/T_s = R_s$ symbols/s without ISI. For ideal Nyquist filtering and zero ISI, the maximum possible symbol transmission rate per hertz- called symbol rate packing- is 2 symbol/s/Hz. It should be noted that the ideal Nyquist filter and the infinite length of its corresponding pulse are not realizable. They can only be approximately realized

A fundamental parameter for communication systems is bandwidth efficiency R_b/B whose units are bits/s/Hz. This represents a measure of data throughput per Hz of bandwidth, and thus measures how efficiently any signaling technique utilizes the bandwidth resource. The Nyquist bandwidth constraint dictates that the theoretical maximum symbol rate packing without ISI is 2 symbols/s/Hz

In M-ary PAM, each symbol is represented by one of M pulse amplitudes where $M = 2^k$. For k=6 or 6-ary PAM, the theoretical maximum efficiency R_b/B without ISI is 12 bits / s Hz.

The more compact we make the signaling spectrum, the higher is the allowable data rate, or the greater is the number of users that can simultaneously be served. Our goal is to reduce the required system bandwidth as much us possible. But what would happen if we try to exceed Nyquist limitation by forcing the system to operate at smaller bandwidths (small filter cut off) than what the constraint dictates? The pulses would become more spread in time, which would degrade the system's error performance due to increased ISI. Although Nyquist filter ensures zero ISI, such a filter has an impulse response tail approaching infinity. Small timing errors can result in ISI, since the zero ISI can be ensured only when the sampling is performed at exactly the correct sampling time. Thus, although a compact spectrum provides optimum bandwidth utilization it is susceptible to ISI degradation induced by timing errors. Therefore we, use a filter

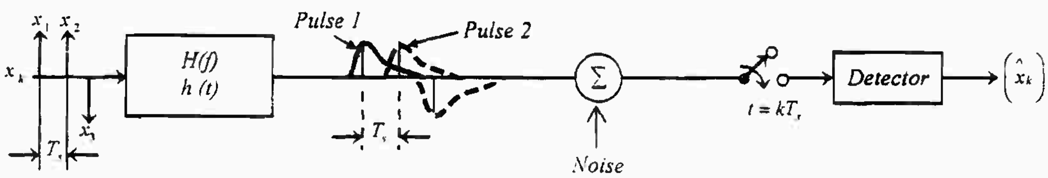


Fig. (2.15) ISI equivalent model of a typical baseband band limited system where the input is approximated to impulses separated by T_s

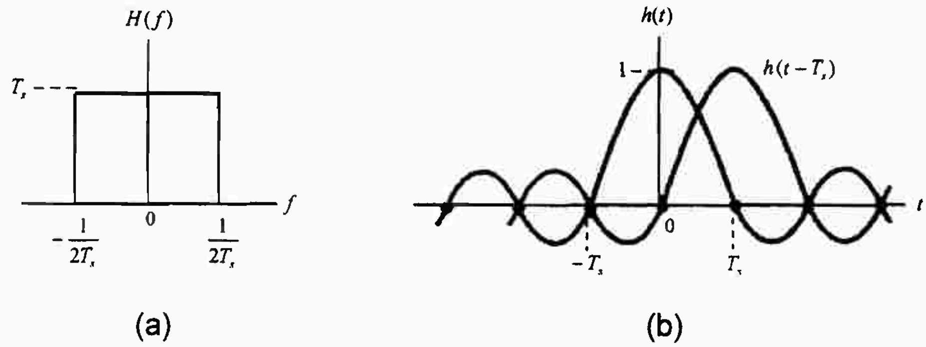


Fig.(2.16). Nyquist constraint for zero ISI
a) rectangular system transfer function
b) received pulse shape $h(t) = \sin t / T_s$

with $H(f)$ having non steep edges to allow for a bandwidth greater than the Nyquist minimum in return for curtailing the pulse tails to reduce ISI even with timing error at the price of increased bandwidth for the same symbol rate or reduced symbol rate for the same bandwidth. One frequently used $H(f)$ transfer function belonging to the Nyquist class (zero ISI at the sampling times) is called the raised cosine filter.

$$H(f) = \begin{cases} 1 & |f| < 2B_o - 3B \\ \cos^2\left(\frac{\pi}{4}\frac{|f| + B + 2B_0}{B - B_o}\right) & 2B_o - B < |f| < B \\ 0 & |f| > B \end{cases} \qquad (2-28)$$

where B is the absolute bandwidth, $B_o = 1/2T_S$ is the minimum Nyquist bandwidth for the rectangular spectrum, and the $6dB$ (or half amplitude point) for the raised cosine spectrum. The difference $B-B_o$ is termed the excess bandwidth, which means the additional bandwidth beyond the Nyquist minimum (i.e. for the rectangular pulse $B-B_o$). The roll off factor is defined to be $r=\frac{B-B_o}{B_o}$, where $0 \le r \le 1$. It represents the excess bandwidth divided by the filter $6dB$ bandwidth, hence called the fractional excess bandwidth. For a given B_o, the roll off r specifies the required excess bandwidth as a fraction of B_o, and characterizes the steepness of the filter roll off. The raised cosine characteristic is illustrated (Fig. 2.17) for roll off values $r=0, 0.5, 1$. The $r=0$ roll off is the Nyquist minimum bandwidth case. For $r=1$, the required excess bandwidth is 100%, and the tails are quite small. A system with such an overall spectral characteristic can provide a symbol rate R_s symbols/s using a bandwidth $R_s Hz$ (twice the Nyquist minimum bandwidth, for fixed R_s thus yielding a symbol rate packing 1 symbol/s/Hz (instead of 2 symbol/s/Hz). The corresponding impulse response of $H(f)$ (eqn. 2 – 28) is given by

$$h(t)=2B_o(\sin 2B_o t)\frac{\cos\left[2\pi(B-B_o)\right]}{1-[4(B-B_o)]^2} \qquad (2-29)$$

We note that the tails have zero value at each pulse sampling time regardless of the roll off value. We also note that the tails are suppressed for $r=1$. However, the raised cosine filter is not psychically realizable. A realizable filter must have an impulse response of finite duration and exhibit a zero output prior to the pulse turn on time. Unrealizable filters are non causal, i.e., the filter impulse response has infinite duration, and the filtered pulse begins at time $t=-\infty$. A pulse shaping filter should satisfy two requirements. It should provide the desired roll off, and it should be realizable, i.e., the impulse response need be truncated to a finite length.

Thus a general relationship between the required bandwidth and the symbol transmission rate involving the filter roll off factor r is given by

$$B=\frac{1}{2}(1+r)R_s \qquad (2-30)$$

For $r=0$, we have Nyquist minimum bandwidth, $B=R_s/2$ and for $r=1$ we have $B=R_s$. We note that in this case we do not have the Nyquist condition, $(R_s=2B)$ for reconstructing the analog signal without aliasing. But in digital communication we may only be interested in detecting the symbols with zero ISI.

(a)

(b)

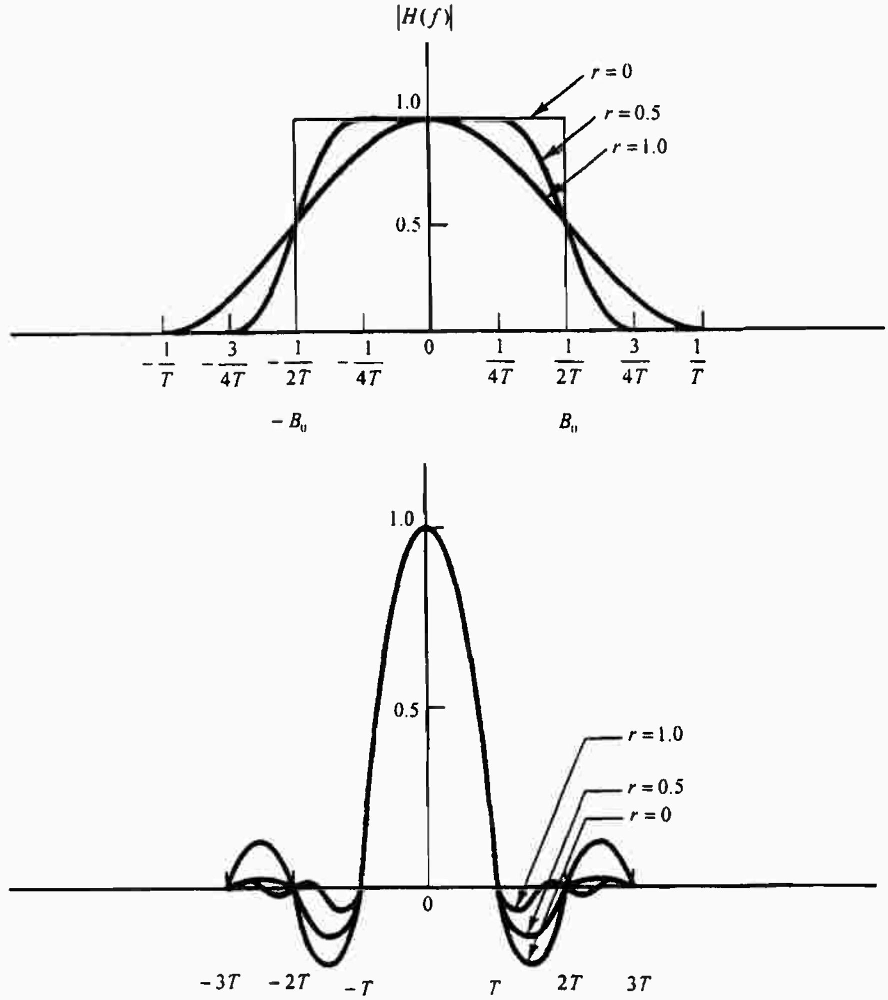


Fig. (2.17) Raised cosine filter

a) system transfer function *b) system impulse response*

Thus

For $r = 0$ $B_1 = R_{s1}/2$ (2 – 31)

For $r = 1$ $B_2 = R_{s2}$ (2 – 32)

For $R_{s1} = R_{s2}$, $B_2 = 2B_1$ (2 – 33)

For $B_1 = B_2$, $R_{s2} = R_{s1}/2$, (2 – 34)

For double side band spectrum Fig. (2.11)

$$B_{DSB} = (1+r)R_s \qquad (2-35)$$

We must note that the raised cosine frequency transfer function describes the composite $H(f)$ that is the full round trip from the transmitter though the channel and through the receiving filter. The receiving filter is where the compensation takes place to bring about zero ISI with an overall transfer function such as the raised cosine. Alternatively, the receiving filter is matched to the transmitting filter, so that each has a transfer function known as the root raised cosine. We note that the longer the filter roll off the shorter will be the pulse tail. The cost is more excess bandwidth. On the other hand, the smaller the filter roll off for a fixed bandwidth the smaller will be the excess bandwidth thereby allowing us to increase the symbol rate or the number of users that can simultaneously use the system. The cost is longer pulse tails, larger pulse amplitude and greater sensitivity to timing errors.

Ex. 2.4

a) Find the minimum required bandwidth for the baseband transmission of a 4 level PAM pulse sequence having a data rate of $R_b = 2400$ bits/s, if the system transfer characteristic consists of a raised cosine spectrum with 100% excess bandwidth $(r=1)$.

b) The same 4-ary PAM is modulated onto a carrier wave so that the baseband spectrum is shifted and centered at frequency f_o. Find the minimum required DSB bandwidth for transmitting the modulated PAM sequence.

Solution

$M = 2^k$, $M = 4$, $k = 2$

Symbol (pulse) rate $R_s = \frac{R_b}{k} = \frac{2400}{2} = 1200$ symbols/s, and $R_s = 1/T_s$

Minimum bandwidth $B = \frac{1}{2}(1+r)R_s = \frac{1}{2}(2)(1200) = 1200\ Hz$

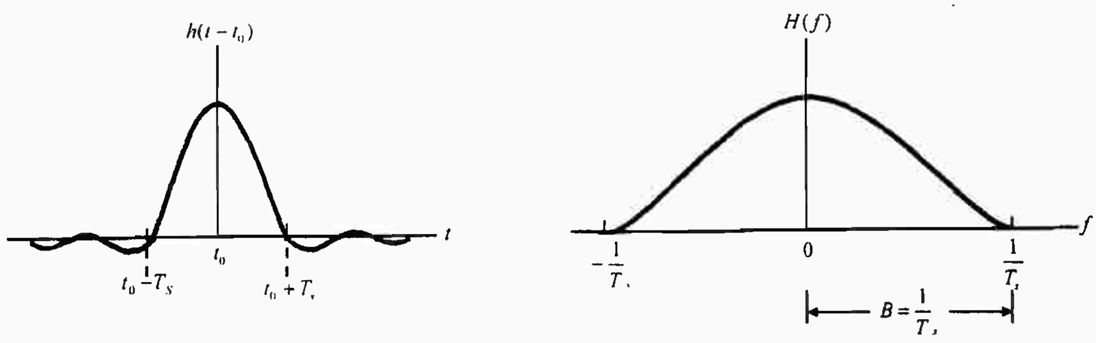
Fig. (2.18), illustrates the baseband PAM received pulse in the time domain which is an approximation to $h(t)$ in eqn. (2–29). The bandwidth B is $1/T_s = R_s$, which is twice the size of Nyquist minimum bandwidth $B_0 = R_s/2$

for DSB, and $r = 1$

$R_s = 1200$ symbols/s

$B_{DSB} = (1+r)R_s = 2(1200) = 2400\ Hz$

$$B_{DSB} = \left(f_o + \frac{1}{T_s}\right) - \left(f_o - \frac{1}{T_s}\right) = \frac{2}{T_s} \qquad (2-37)$$

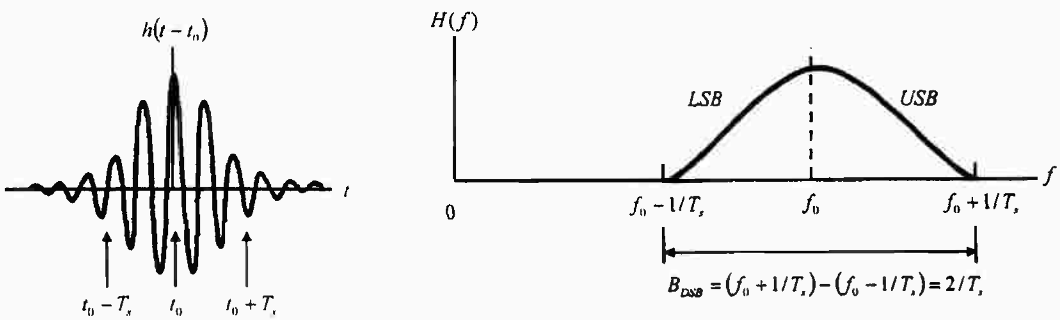


(a) (b)

Fig. 2.18 Impulse transfer response and transfer function

a) shaped pulse impulse response function

b) baseband raised cosine spectrum



(a) (b)

Fig. 2.19 Modulated PAM received pulse.

a) modulated shaped pulse

b) DSB modulated raised cosine spectrum

Ex. 2.5

Compare the system bandwidth for a terrestrial $3kHz$ analog telephone voice channel with that of a digital one for which the voice is formatted as PCM bit stream when the sampling rate is 8000 samples/s and each voice sample is quantized to one of 256 levels. The bit stream is transmitted using a PCM waveform and received with zero ISI.

Solution

Each sample is sent as 256-ary PAM. The word size is 8 bits (eqn. 2 – 4), we require $B \geq R_s / 2$, where the equality sign holds only for ideal Nyquist filtering.

Thus, $L = 2^n = 256$, $n = 8$, $R_b = nR_s = 8R_s$

$$B_{PCM} \geq (\log_2 L)\frac{R_s}{2} \qquad (2-38)$$

$$\geq \frac{1}{2}(8)(8000) \text{ symbols/s} = 32\, kH_z$$

The $3\, kHz$ analog voice channel will generally require $4\, kHz$ of bandwidth separation between channels (guard bands), while the PCM format using 8 bit quantization and binary signaling with PCM requires at least 8 times the bandwidth required for the analog channel. The gain is all the advantages of digital telephony.

Problems

1. The information in an analog waveform whose maximum frequency $f_m = 4\ kHz$ is to be transmitted using 16 level PAM. The quantization distortion must not exceed $\pm 1\%$ of the peak to peak analog signal?
   a) What is the minimum number of bits per sample or bits per PCM word?
   b) What is the minimum required sampling rate and bit rate?
   c) What is the 16-ary PAM symbol rate?

2. A signal with frequency range $40\ Hz$ to $4\ kHz$ is limited to peak to peak amplitude variation of $12\ V$. It is sampled at 6000 samples/s and the samples are quantized to 64 evenly spaced levels. Calculate and compare the bandwidths and ratios of peak signal power to rms quantization noise if the quantized samples are transmitted
   a) as binary pulses b) as 4 level pulses

3. For a CD digital audio system, an analog signal is digitized so that the ratio of the peak signal power to the peak quantization noise power is at least $96\ dB$. The sampling rate is $44.1k$ samples/s.
   a) How many quantization levels of the analog signal are needed?
   b) How many bits / sample are needed?
   c) What is the data rate?

4. Assume the quantization distortion of an audio source is $\pm 1\%$ of the peak to peak analog signal voltage. The audio signal bandwidth and the allowable transmission bandwidth are each $4\ kHz$ and samples are taken at the Nyquist rate. Calculate the bandwidth efficiency.

5. Repeat the problem above if the signal bandwidth is $20\ kHz$ (Hi Fi), and the transmission bandwidth is still $4\ kHz$.

6. Find the minimum required bandwidth for the baseband transmission of a 4 level PAM pulse sequence having a data rate of $R_b = 2400$ bits/s if the system transfer characteristic is raised cosine with 100% excess bandwidth $(r = 1)$ i.e. twice the Nyquist minimum bandwidth.

7. Repeat Prob. 6 if the symbol rate is doubled. Comment.

8. Repeat Prob. 6 if the available bandwidth is reduced to half. Comment.

9. In Prob. 8 the baseband is shifted and centered at frequency f_0, find the minimum required DSB bandwidth.

10. Compare the system bandwidth requirements for a $4\ kHz$ telephone voice channel with that of a digital channel in which the sampling rate is 6000 samples/s and each sample is quantized to one of 64 levels given that PCM with zero ISI is used.

References

1. "Digital Communications," B. Sklar, Prentice Hall, Upper Saddle River, N.J., 2001.
2. "Telecommunications Demystified," C. Nassar, LLH Technology Publishing, Eagle Rock, Virginia, 2001.
3. "Communication Engineering Principles," I. Otung, Palgrave, N.Y., 2001.
4. "Digital Communications," I. Glover, P. Grant, Prentice Hall Europe. Hemel Hempstead, Hertfordshire, GB, 1998.
5. "Digital Communications," J. Proakis, 4^{th} ed, McGraw Hill, N.Y, 2001.